SUPPLEMENTARY: PROOFS FOR THEOREM 1,2 AND 3

A. Some Lemmas

In order to prove Theorem 3, we need to introduce the following Lemmas $1 \sim 6$.

Lemma 1. *(Lemma 1 in [25]): For nonnegative numbers a, b, c, d, x, y,*

$$(ax+by)^2+(cx+dy)^2 \le \left(\sqrt{a^2+c^2}x+(b+d)y\right)^2. \quad (16)$$

Lemma 2. *Consider the general CS model $\mathbf{Y}=\mathbf{A}\mathbf{X}+\mathbf{E}$, supp$(\mathbf{X})=T$ and $|T|=s$. Suppose $S, T_0 \subseteq \{1,2,...,n\}$, $|S|=t$. $\mathbf{W}_{T_0}$ is constructed with diagonal entries indexed by T_0 being $\omega \ge 0$. Let $\tilde{\mathbf{X}}$ be the solution of the least squares problem $\arg\min_{\mathbf{Z}} \{\|\mathbf{Y}-\mathbf{A}\mathbf{Z}\|_2, \text{supp}(\mathbf{Z}) \subseteq S\}$. Let $s_1 := \min_{i,k}|\Lambda_{S_i}^k \cap \text{supp}(\mathbf{X}_{S_i})|$, $s_2 := \min_{i,k}|\Lambda_{S_i}^k \cap \text{supp}(\mathbf{X}_{S_i}) \cap \tilde{S}_{S_i}^k|$, $s_3 := \min_{i,k}|\Lambda_{S_i}^k \cap \text{supp}(\mathbf{X}_{S_i}) \cap \tilde{S}_{S_i}^k \setminus S_{S_i}^k|$ and $\eta = \min\limits_{i\in\{1,2,...,r\}} \min\limits_{j\in\{1,2,...,n\}} \|(\mathbf{X}_{S_i})_{j,:}\|_2 / \|\mathbf{X}_{S_i}\|_F$. Let $\nu_1 = (1+\omega)/(1+\eta\omega\sqrt{s_2})$ and $\nu_2=(1+\omega)/(1+\eta\omega\sqrt{s_3})$, if $\delta_{3s}<1$, then*

$$\left\|\mathbf{W}_{T_0}(\mathbf{X}-\tilde{\mathbf{X}})_S\right\|_F \le \omega\delta_{s+t}\left\|\mathbf{X}-\tilde{\mathbf{X}}\right\|_F + \omega\sqrt{1+\delta_t}\,\|\mathbf{E}\|_F \quad (17)$$

and

$$\left\|\mathbf{X}-\tilde{\mathbf{X}}\right\|_F \le \sqrt{\frac{1}{1-\delta_{s+t}^2}}\,\|(\mathbf{X})_T\|_F + \frac{\sqrt{1+\delta_t}}{1-\delta_{s+t}}\|\mathbf{E}\|_F \quad (18)$$

If $t>s$, define $T_\triangledown$ as the row-indices of the smallest $t-s$ magnitude entries of $\tilde{\mathbf{X}}$ in S, we have

$$\|\mathbf{X}_{T_\triangledown}\|_F \le \sqrt{2}\nu_2\delta_{s+t}\left\|\mathbf{X}-\tilde{\mathbf{X}}\right\|_F + \nu_2\sqrt{2(1+\delta_t)}\,\|\mathbf{E}\|_F. \quad (19)$$

- First, we give a upper bound of $\|\mathbf{X}_{T_\triangledown}\|_F$, by Lemma 6, let $\mathbf{Z}=(\mathbf{W}_{T_0}\mathbf{X}-\mathbf{W}_{T_0}\tilde{\mathbf{X}})_S$, we have

$$\left\langle \mathbf{W}_{T_0}(\mathbf{X}-\tilde{\mathbf{X}}), \mathbf{A}^H\mathbf{A}(\mathbf{W}_{T_0}\mathbf{X}-\mathbf{W}_{T_0}\tilde{\mathbf{X}})_S\right\rangle + \left\langle \mathbf{W}_{T_0}\mathbf{E}, \mathbf{A}(\mathbf{W}_{T_0}\mathbf{X}-\mathbf{W}_{T_0}\tilde{\mathbf{X}})_S\right\rangle = 0 \quad (20)$$

  Noticing that supp$(\tilde{\mathbf{X}}) \subseteq S$, we have

$$\begin{aligned}
&\left\|(\mathbf{W}_{T_0}\mathbf{X}-\mathbf{W}_{T_0}\tilde{\mathbf{X}})_S\right\|_F^2 \\
=\ &\left\langle \mathbf{W}_{T_0}(\mathbf{X}-\tilde{\mathbf{X}}), (\mathbf{W}_{T_0}\mathbf{X}-\mathbf{W}_{T_0}\tilde{\mathbf{X}})_S\right\rangle \\
\overset{(20)}{=}\ &\left\langle \mathbf{W}_{T_0}(\mathbf{X}-\tilde{\mathbf{X}}), (\mathbf{I}_L-\mathbf{A}^H\mathbf{A})(\mathbf{X}-\tilde{\mathbf{X}})_S\right\rangle \\
-\ &\left\langle \mathbf{W}_{T_0}\mathbf{E}, \mathbf{A}(\mathbf{W}_{T_0}\mathbf{X}-\mathbf{W}_{T_0}\tilde{\mathbf{X}})_S\right\rangle \\
\overset{(7)}{\le}\ &\omega\delta_{s+t}\left\|(\mathbf{W}_{T_0}\mathbf{X}-\mathbf{W}_{T_0}\tilde{\mathbf{X}})_S\right\|_F\left\|\mathbf{X}-\tilde{\mathbf{X}}\right\|_F \\
&+\omega\|\mathbf{E}\|_F\sqrt{1+\delta_t}\left\|\mathbf{W}_{T_0}(\mathbf{X}-\tilde{\mathbf{X}})_S\right\|_F \quad (21)
\end{aligned}$$

  Divide both sides by $\left\|(\mathbf{W}_{T_0}\mathbf{X}-\mathbf{W}_{T_0}\tilde{\mathbf{X}})_S\right\|_F$ to obtain (17).

- By expanding Lemma 2 in [25] to MMV model, we could get a relationship between $\left\|\mathbf{X}-\tilde{\mathbf{X}}\right\|_F$ and $\|\mathbf{X}_{\overline{S}}\|_F$, we have

$$\left\|\mathbf{X}-\tilde{\mathbf{X}}\right\|_F \le \sqrt{\frac{1}{1-\delta_{s+t}^2}\|\mathbf{X}_{\overline{S}}\|_F^2} + \frac{\sqrt{1+\delta_t}}{1-\delta_{s+t}}\|\mathbf{E}\|_F \quad (22)$$

- Finally, we established the relationship between $\mathbf{X}_{T_\triangledown}$ and $\mathbf{X}-\tilde{\mathbf{X}}$. There exist a subset $S' \subseteq S$ and $S' \cap T = \emptyset$. Since $T_\triangledown$ is defined by the set of indices of the $t-s$ smallest row entries of $\tilde{\mathbf{X}}$, we can conclude that

$$\begin{aligned}
&\left\|\tilde{\mathbf{X}}_{T_\triangledown}\right\|_F + \left\|\mathbf{W}_{T_0}\tilde{\mathbf{X}}_{T_\triangledown}\right\|_F \\
&\le \left\|\tilde{\mathbf{X}}_{S'}\right\|_F + \left\|\mathbf{W}_{T_0}\tilde{\mathbf{X}}_{S'}\right\|_F \quad (23)
\end{aligned}$$

By eliminating the contribution from $T_\triangledown \cap S'$, and noticing that $S' \cap T = \emptyset$ we have

$$\begin{aligned}
&\left\|\tilde{\mathbf{X}}_{T_\triangledown\setminus S'}\right\|_F + \left\|\mathbf{W}_{T_0}\tilde{\mathbf{X}}_{T_\triangledown\setminus S'}\right\|_F \\
\le\ &\left\|(\tilde{\mathbf{X}}-\mathbf{X})_{S'\setminus T_\triangledown}\right\|_F \\
+\ &\left\|\mathbf{W}_{T_0}(\tilde{\mathbf{X}}-\mathbf{X})_{S'\setminus T_\triangledown}\right\|_F \quad (24)
\end{aligned}$$

For the lefthand side of (24), we have

$$\begin{aligned}
&\left\|\tilde{\mathbf{X}}_{T_\triangledown\setminus S'}\right\|_F + \left\|\mathbf{W}_{T_0}\tilde{\mathbf{X}}_{T_\triangledown\setminus S'}\right\|_F \\
=\ &\left\|(\tilde{\mathbf{X}}-\mathbf{X}+\mathbf{X})_{T_\triangledown\setminus S'}\right\|_F \\
&+\left\|(\mathbf{W}_{T_0}(\tilde{\mathbf{X}}-\mathbf{X})+\mathbf{W}_{T_0}\mathbf{X})_{T_\triangledown\setminus S'}\right\|_F \\
\ge\ &\|\mathbf{X}_{T_\triangledown}\|_F + \|\mathbf{W}_{T_0}\mathbf{X}_{T_\triangledown}\|_F \quad (25)\\
-\ &\left\|(\tilde{\mathbf{X}}-\mathbf{X})_{T_\triangledown\setminus S'}\right\|_F \\
-\ &\left\|\mathbf{W}_{T_0}(\tilde{\mathbf{X}}-\mathbf{X})_{T_\triangledown\setminus S'}\right\|_F \quad (26)
\end{aligned}$$

Combining (26) and (24), and noticing that

$$(T_\triangledown \setminus S') \cap (S' \setminus T_\triangledown) = \emptyset \quad (27)$$

$$(T_\triangledown \setminus S') \cup (S' \setminus T_\triangledown) \subseteq T \quad (28)$$

we have

$$\begin{aligned}
&\|\mathbf{X}_{T_\triangledown}\|_F + \|\mathbf{W}_{T_0}\mathbf{X}_{T_\triangledown}\|_F \\
&\le \left\|(\tilde{\mathbf{X}}-\mathbf{X})_{T_\triangledown\setminus S'}\right\|_F + \left\|(\mathbf{W}_{T_0}(\tilde{\mathbf{X}}-\mathbf{X}))_{T_\triangledown\setminus S'}\right\|_F \\
&+\left\|(\tilde{\mathbf{X}}-\mathbf{X})_{S'\setminus T_\triangledown}\right\|_F + \left\|(\mathbf{W}_{T_0}(\tilde{\mathbf{X}}-\mathbf{X}))_{S'\setminus T_\triangledown}\right\|_F \\
&\le \sqrt{2}\left\|(\tilde{\mathbf{X}}-\mathbf{X})_{(S'\setminus T_\triangledown)\cup(T_\triangledown\setminus S')}\right\|_F \\
&+\sqrt{2}\left\|(\mathbf{W}_{T_0}(\tilde{\mathbf{X}}-\mathbf{X}))_{(S'\setminus T_\triangledown)\cup(T_\triangledown\setminus S')}\right\|_F \\
&\le \sqrt{2}\left\|(\tilde{\mathbf{X}}-\mathbf{X})_S\right\|_F + \sqrt{2}\left\|(\mathbf{W}_{T_0}(\tilde{\mathbf{X}}-\mathbf{X}))_S\right\|_F \\
&\overset{(17)}{\le} \sqrt{2}(1+\omega)\delta_{s+t}\left\|\mathbf{X}-\tilde{\mathbf{X}}\right\|_F \\
&+(1+\omega)\sqrt{2(1+\delta_t)}\,\|\mathbf{E}\|_F \quad (29)
\end{aligned}$$

We can obtain the relationship between $\left\|\mathbf{W}_{T_0}\mathbf{X}_{T_\triangledown}\right\|_F$ and $\left\|\mathbf{X}_{T_\triangledown}\right\|_F$

$$\eta\omega\sqrt{s_3}\left\|\mathbf{X}_{T_\triangledown}\right\|_F \le \left\|\mathbf{W}_{T_0}\mathbf{X}_{T_\triangledown}\right\|_F \tag{30}$$

Noting the definition of ν_2, combining (29) and (30), we have

$$\begin{aligned}\left\|\mathbf{X}_{T_\triangledown}\right\|_F \le & \frac{\sqrt{2}(1+\omega)\delta_{s+t}}{1+\eta\omega\sqrt{s_3}}\left\|\mathbf{X}-\tilde{\mathbf{X}}\right\|_F \\ & +\frac{(1+\omega)\sqrt{2(1+\delta_t)}}{1+\eta\omega\sqrt{s_3}}\left\|\mathbf{E}\right\|_F\end{aligned} \tag{31}$$

We have completed the proof of Lemma 2.

When we consider the atom selection strategy of $\left\|\tilde{\mathbf{X}}_{T_\triangledown}+\mathbf{W}_{T_0}\tilde{\mathbf{X}}_{T_\triangledown}\right\|_F \le \left\|\tilde{\mathbf{X}}_{S'}+\mathbf{W}_{T_0}\tilde{\mathbf{X}}_{S'}\right\|_F$, we can also obtain another upper bound for $\left\|\mathbf{X}_{T_\triangledown}\right\|_F$ in (19). In this case, we should allocate $2\left\|\mathbf{X}_{T_\triangledown}\right\|_F$ to the left hand side of (25), we have

$$\begin{aligned}\left\|\mathbf{X}_{T_\triangledown}\right\|_F \le & \sqrt{2}\nu_3\delta_{s+t}\left\|\mathbf{X}-\tilde{\mathbf{X}}\right\|_F \\ & +\nu_4\sqrt{2\left(1+\delta_t\right)}\left\|\mathbf{E}\right\|_F.\end{aligned} \tag{32}$$

where $\nu_3 = (1-\omega+\omega\delta_{s+t}+\delta_{s+t})/(2\delta_{s+t})$ and $\nu_4 = (1+\omega)/(2\delta_{s+t})$.

Lemma 3. *In steps 4 and 5 of SI-SSP, we have*

$$\left\|\mathbf{X}_{\overline{\tilde{S}^k}}\right\|_F \le \sqrt{2}\varphi_1\delta_{3s}\left\|\mathbf{X}-\mathbf{X}^{k-1}\right\|_F + \varphi_1\sqrt{2(1+\delta_{3s})}\left\|\mathbf{E}\right\|_F. \tag{33}$$

Proof: From step 5 of IS-SSP, we have

$$\mathbf{X}^k_{S_i} = \underset{\boldsymbol{\Theta}:\text{supp}(\boldsymbol{\Theta})=S^k_{S_i}}{\arg\min}\left\|\mathbf{Y}_{S_i}-\mathbf{A}\boldsymbol{\Theta}\right\|_F \tag{34}$$

From the step 4 of SI-SSP, let $\mathbf{X}^k = [\mathbf{X}^k_{S_1},\cdots,\mathbf{X}^k_{S_r}]$, we have the following conclusion

$$\begin{aligned}&\left\|(\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_T\right\|_F \\ &+\left\|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_T\right\|_F \\ &\le \left\|(\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{\triangle S}\right\|_F \\ &+\left\|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{\triangle S}\right\|_F\end{aligned} \tag{35}$$

By removing the same coordinates $T\cap\triangle S$, we can get

$$\begin{aligned}&\left\|(\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{T\backslash\triangle S}\right\|_F \\ &+\left\|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{T\backslash\triangle S}\right\|_F \\ &\le \left\|(\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{\triangle S\backslash T}\right\|_F \\ &+\left\|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{\triangle S\backslash T}\right\|_F\end{aligned} \tag{36}$$

Because $\text{supp}(\mathbf{X}) = T$ and $\text{supp}(\mathbf{X}^{k-1}) = S^{k-1}$

$$(\mathbf{X}-\mathbf{X}^{k-1})_{\triangle S\backslash(T\cup S^{k-1})} = 0 \tag{37}$$

For the right-hand of (36), we have

$$\begin{aligned}&\left\|(\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{\triangle S\backslash T}\right\|_F \\ &+\left\|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{\triangle S\backslash T}\right\|_F \\ &= \left\|(\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{\Delta S\backslash(T\cup S^{k-1})}\right\|_F \\ &+\left\|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{\Delta S\backslash(T\cup S^{k-1})}\right\|_F \\ &= \left\|(\mathbf{A}^H(\mathbf{A}\mathbf{X}+\mathbf{E}-\mathbf{A}\mathbf{X}^{k-1}))_{\Delta S\backslash(T\cup S^{k-1})}\right\|_F \\ &+\left\|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{A}\mathbf{X}+\mathbf{E}-\mathbf{A}\mathbf{X}^{k-1}))_{\Delta S\backslash(T\cup S^{k-1})}\right\|_F \\ &\le \left\|((\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{\Delta S\backslash T}\right\|_F \\ &+\left\|(\mathbf{A}^H\mathbf{E})_{\Delta S\backslash T}\right\|_F \\ &+\left\|(\mathbf{W}_{T_0}(\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{\Delta S\backslash T}\right\|_F \\ &+\left\|(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{E})_{\Delta S\backslash T}\right\|_F\end{aligned} \tag{38}$$

Note that $\tilde{S}^k = S^{k-1}\cup\Delta S$, we have

$$(\mathbf{X}-\mathbf{X}^{k-1})_{T\backslash\tilde{S}^k} = \mathbf{X}_{\overline{\tilde{S}^k}} \tag{39}$$

For the left-side of (36),we have

$$\begin{aligned}&\left\|(\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{T\backslash\Delta S}\right\|_F \\ &+\left\|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{T\backslash\Delta S}\right\|_F \\ &= \left\|(\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{T\backslash(\Delta S\cup S^{k-1})}\right\|_F \\ &+\left\|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{T\backslash(\Delta S\cup S^{k-1})}\right\|_F \\ &= \left\|(\mathbf{A}^H(\mathbf{A}\mathbf{X}+\mathbf{E}-\mathbf{A}\mathbf{X}^{k-1}))_{T\backslash(\Delta S\cup S^{k-1})}\right\|_F \\ &+\left\|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{A}\mathbf{X}+\mathbf{E}-\mathbf{A}\mathbf{X}^{k-1}))_{T\backslash(\Delta S\cup S^{k-1})}\right\|_F \\ &= \left\|((\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1})\right. \\ &\left.+\mathbf{A}^H\mathbf{E}+\mathbf{X})_{\Delta S\backslash(T\cup S^{k-1})}\right\| \\ &+\left\|\mathbf{W}_{T_0}(\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1})\right. \\ &\left.+\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{E}+\mathbf{W}_{T_0}\mathbf{X})_{\Delta S\backslash(T\cup S^{k-1})}\right\|_F \\ &\ge \left\|\mathbf{X}_{\overline{\tilde{S}^k}}\right\|_F+\left\|(\mathbf{W}_{T_0}\mathbf{X})_{\overline{\tilde{S}^k}}\right\|_F \\ &-\left\|((\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{\overline{\tilde{S}^k}}\right\|_F \qquad (40) \\ &-\left\|(\mathbf{W}_{T_0}(\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{\overline{\tilde{S}^k}}\right\|_F \\ &-\left\|(\mathbf{A}^H\mathbf{E})_{\overline{\tilde{S}^k}}\right\|_F-\left\|(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{E})_{\overline{\tilde{S}^k}}\right\|_F\end{aligned} \tag{41}$$

Combining (42) and (41), we have

$$\begin{aligned}&\left\|\mathbf{X}_{\overline{\tilde{S}^k}}\right\|_F+\left\|(\mathbf{W}_{T_0}\mathbf{X})_{\overline{\tilde{S}^k}}\right\|_F \\ &\le \left\|((\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{T\backslash\tilde{S}^k}\right\|_F \\ &+\left\|(\mathbf{W}_{T_0}(\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{T\backslash\tilde{S}^k}\right\|_F \\ &+\left\|(\mathbf{W}_{T_0}(\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{\Delta S\backslash T}\right\|_F \\ &+\left\|((\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{\Delta S\backslash T}\right\|_F \\ &+\left\|(\mathbf{A}^H\mathbf{E})_{T\backslash\tilde{S}^k}\right\|_F+\left\|(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{E})_{T\backslash\tilde{S}^k}\right\|_F \\ &+\left\|(\mathbf{A}^H\mathbf{E})_{\Delta S\backslash T}\right\|_F+\left\|(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{E})_{\Delta S\backslash T}\right\|_F \\ &\le \sqrt{2}\left\|((\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{T\cup\Delta S}\right\|_F \\ &+\sqrt{2}\left\|(\mathbf{W}_{T_0}(\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{T\cup\Delta S}\right\|_F \\ &+\sqrt{2}\left\|(\mathbf{A}^H\mathbf{E})_{T\cup\Delta S}\right\|_F+\sqrt{2}\left\|(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{E})_{T\cup\Delta S}\right\|_F \\ &\le \sqrt{2}(1+\omega)\delta_{3s}\left\|\mathbf{X}-\mathbf{X}^{k-1}\right\|_F \\ &+(1+\omega)\sqrt{2(1+\delta_{3s})}\left\|\mathbf{E}\right\|_F\end{aligned} \tag{42}$$

We can obtain the relationship between $\left\|\mathbf{X}_{\overline{\tilde{S}^k}}\right\|_F$ $\left\|(\mathbf{W}_{T_0}\mathbf{X})_{\overline{\tilde{S}^k}}\right\|_F$

$$\eta\omega\sqrt{s_2}\left\|\mathbf{X}_{\overline{\tilde{S}^k}}\right\|_F \leq \left\|(\mathbf{W}_{T_0}\mathbf{X})_{\overline{\tilde{S}^k}}\right\|_F \tag{43}$$

Noting the definition of ν_1, combining (42) and (43), we have

$$\begin{aligned}\left\|\mathbf{X}_{\overline{\tilde{S}^k}}\right\|_F \leq& \frac{\sqrt{2}(1+\omega)\delta_{3s}}{1+\eta\omega\sqrt{s_2}}\left\|\mathbf{X}-\tilde{\mathbf{X}}\right\|_F \\ &+\frac{(1+\omega)\sqrt{2(1+\delta_{2s})}}{1+\eta\omega\sqrt{s_2}}\|\mathbf{E}\|_F\end{aligned} \tag{44}$$

We have completed the proof of Lemma 3.

When we consider the atom selection strategy in select step of

$$\begin{aligned}&\left\|((\mathbf{I}_L+\mathbf{W}_{T_0})\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_T\right\|_F \\ \leq& \left\|((\mathbf{I}_L+\mathbf{W}_{T_0})\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{\triangle S}\right\|_F\end{aligned} \tag{45}$$

We can also obtain another upper bound for $\left\|\mathbf{X}_{\overline{\tilde{S}^k}}\right\|_F$ in (33). In this case, we should allocate $2\left\|\mathbf{X}_{\overline{\tilde{S}^k}}\right\|_F$ to the left hand side of (40), we have

$$\begin{aligned}\left\|\mathbf{X}_{\overline{\tilde{S}^k}}\right\|_F \leq& \sqrt{2}\nu_4\delta_{3s}\left\|\mathbf{X}-\mathbf{X}^{k-1}\right\|_F \\ &+\nu_4\sqrt{2\left(1+\delta_{3s}\right)}\|\mathbf{E}\|_F\end{aligned} \tag{46}$$

where $\nu_4 = (1-\omega+\omega\delta_{3s}+\delta_{3s})/(2\delta_{3s})$ and $\nu_4 = (1+\omega)/(2\delta_{3s})$. Based on conclusions (32) and (46), we know that the sensing matrix $\mathbf{A}$ obeys the RIP with

$$\delta_{3s} \leq \sqrt{\frac{\nu_3\sqrt{\nu_3^2+4\nu_4^2}-\nu_3^2-1}{4\nu_3^2\nu_4^2-2\nu_3^2-1}} \tag{47}$$

Lemma 4. *Let $T_0 \subseteq \{1,2,...,n\}$, for two vectors $\mathbf{u},\mathbf{v} \in R^n$, if $|supp(\mathbf{u}) \cup supp(\mathbf{v})| \leq t$,*

$$\left|\left\langle\mathbf{u},(\mathbf{W}_{T_0}-\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{A})\mathbf{v}\right\rangle\right| \leq \omega\delta_t\|\mathbf{u}\|\|\mathbf{v}\|; \tag{48}$$

moreover if $U \subseteq \{1,2,...,n\}$ and $|U \cup supp(\mathbf{v})| \leq t$, then

$$\left|(\mathbf{W}_{T_0}-\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{A})\mathbf{v}\right| \leq \omega\delta_t\|\mathbf{v}\|. \tag{49}$$

Proof: the RIC δ_s can be expressed as [18]

$$\delta_s = \max_{S\subseteq\{1,2,...,N\},|S|\leq s}\|\mathbf{A}_S^*\mathbf{A}_S-\mathbf{I}\|_{2\to 2}, \tag{50}$$

where

$$\|\mathbf{A}_S^*\mathbf{A}_S-\mathbf{I}\|_{2\to 2} = \sup_{\mathbf{a}\in\mathbb{R}^{|S|}\setminus\{\mathbf{0}\}}\frac{\|(\mathbf{A}_S^*\mathbf{A}_S-\mathbf{I})\mathbf{a}\|_2}{\|\mathbf{a}\|_2}. \tag{51}$$

Let $S = \mathrm{supp}(\mathbf{u}) \cup \mathrm{supp}(\mathbf{v})$. Then $|S| \leq t$. Let $\mathbf{u}_{|S}, \mathbf{v}_{|S}$ denote respectively the S-dimensional sub-vectors of $\mathbf{u}$ and $\mathbf{v}$ obtained by only keeping the components indexed by S. We have

$$\begin{aligned}&\left|\left\langle\mathbf{u},\left(\mathbf{W}_{T_0}-\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{A}\right)\mathbf{v}\right\rangle\right| \\ =&\left|\langle\mathbf{W}_{T_0}\mathbf{u},\mathbf{v}\rangle-\langle\mathbf{A}\mathbf{W}_{T_0}\mathbf{u},\mathbf{A}\mathbf{v}\rangle\right| \\ =&\left|\left\langle\mathbf{W}_{T_0}\mathbf{u}_{|T},\left(\mathbf{I}_L-\mathbf{A}_T^H\mathbf{A}_T\right)\mathbf{v}_{|T}\right\rangle\right| \\ \leq&\left\|\mathbf{W}_{T_0}\mathbf{u}_{|T}\right\|_2\left\|\left(\mathbf{I}_L-\mathbf{A}_T^H\mathbf{A}_T\right)\mathbf{v}_{|T}\right\|_2 \\ \overset{(51)}{\leq}&\left\|\mathbf{W}_{T_0}\mathbf{u}_{|T}\right\|_2\left\|\mathbf{I}_L-\mathbf{A}_T^H\mathbf{A}_T\right\|_{2\to 2}\left\|\mathbf{v}_{|T}\right\|_2 \\ \overset{(50)}{\leq}&\omega\delta_t\left\|\mathbf{u}_{|T}\right\|_2\left\|\mathbf{v}_{|T}\right\|_2 \\ =&\omega\delta_t\|\mathbf{u}\|_2\|\mathbf{v}\|_2,\end{aligned} \tag{52}$$

Moreover, we have

$$\begin{aligned}&\left\|\left(\left(\mathbf{W}_{T_0}-\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{A}\right)\mathbf{v}\right)_U\right\|_2^2 \\ =&\left\langle\left(\left(\mathbf{W}_{T_0}-\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{A}\right)\mathbf{v}\right)_U\right. \\ &\left.,\left(\mathbf{W}_{T_0}-\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{A}\right)\mathbf{v}\right\rangle \\ \overset{(48)}{\leq}&\delta_t\left\|\left(\left(\mathbf{W}_{T_0}-\mathbf{A}^H\mathbf{A}\right)\mathbf{v}\right)_U\right\|_2\|\mathbf{v}\|_2,\end{aligned} \tag{53}$$

which completes the proof of Lemma 4.

Lemma 5. *For SMV model $\mathbf{y} = \Phi\mathbf{x}+\mathbf{e}$, let $T_0 \subseteq \{1,2,...,n\}$, let $U \subseteq \{1,2,...,n\}$ and $|U \cap T_0| \leq u$, we have*

$$\left\|(\mathbf{W}_{T_0}\mathbf{A}^He)_U\right\|_2 \leq \omega\delta_u\|e\|_2 \tag{54}$$

Proof: The lemma easily follows from the fact that

$$\begin{aligned}&\left\|\left(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{e}\right)_U\right\|_2^2 \\ =&\left\langle\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{e},\left(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{e}\right)_U\right\rangle \\ =&\left\langle\mathbf{e},\mathbf{W}_{T_0}\mathbf{A}\left(\left(\mathbf{A}^H\mathbf{e}\right)_U\right)\right\rangle \\ \leq&\|\mathbf{e}\|_2\left\|\mathbf{W}_{T_0}\mathbf{A}\left(\left(\mathbf{A}^H\mathbf{e}\right)_U\right)\right\|_2 \\ \overset{(3)}{\leq}&\|\mathbf{e}'\|_2\,\omega\sqrt{1+\delta_u}\left\|\left(\mathbf{A}^H\mathbf{e}\right)_U\right\|_2.\end{aligned} \tag{55}$$

Lemma 6. *Consider MMV model $\mathbf{Y} = \mathbf{A}\mathbf{X}+\mathbf{E}$, let $\tilde{\mathbf{X}}$ be the solution of the least squares problem $\arg\min_{\mathbf{Z}}\{\|\mathbf{Y}-\mathbf{A}\mathbf{Z}\|_F, supp(\mathbf{Z}) \subseteq S\}$, then*

$$\left\langle\mathbf{W}_{T_0}\mathbf{X}-\mathbf{W}_{T_0}\tilde{\mathbf{X}},\mathbf{A}^H\mathbf{A}\mathbf{Z}\right\rangle+\omega\langle\mathbf{E},\mathbf{A}\mathbf{Z}\rangle = 0 \tag{56}$$

Proof: Due to the orthogonality, the residue $\mathbf{Y}-\mathbf{A}\tilde{\mathbf{X}}$ is orthogonal to the space $\mathbf{A}\mathbf{Z}, \mathrm{supp}(\mathbf{Z}) \subseteq S$. This means that for all $\mathbf{Z} \in \mathcal{C}$ with $\mathrm{supp}(\mathbf{Z}) \subseteq S$,

$$\left\langle\mathbf{A}(\mathbf{Y}-\mathbf{A}\tilde{\mathbf{X}}),\mathbf{Z}\right\rangle = 0 \tag{57}$$

then, let $\tilde{X}'$ be the solution of the least squares problem $\arg\min_{\mathbf{Z}}\{\|\mathbf{Y}'-\mathbf{A}\mathbf{Z}\|_F, \mathrm{supp}(\mathbf{Z}) \subseteq S\}$, where $\mathbf{Y}' = \frac{\mathbf{A}\mathbf{W}_{T_0}\mathbf{X}_{T_0}}{\omega}+\mathbf{E}$ we have

$$\tilde{X}' = \frac{\mathbf{W}_{T_0}\tilde{X}}{\omega} \tag{58}$$

Then, by (57), we have

$$\begin{aligned}0 &= \left\langle\frac{\mathbf{A}\mathbf{W}_{T_0}\mathbf{X}_{T_0}}{\omega}+\mathbf{E}-\mathbf{A}\frac{\mathbf{W}_{T_0}\tilde{X}}{\omega},\mathbf{A}\mathbf{Z}\right\rangle \\ &= \left\langle\mathbf{W}_{T_0}\mathbf{X}-\mathbf{W}_{T_0}\tilde{\mathbf{X}},\mathbf{A}^H\mathbf{A}\mathbf{Z}\right\rangle+\omega\langle\mathbf{E},\mathbf{A}\mathbf{Z}\rangle.\end{aligned} \tag{59}$$

B. Proof of Theorem 1

Firstly, we assume that a multi-coset sampler has p channels, and the sampling rate of each channel is determined by the sampled signal sequence, which is in the form of:

$$x_{c_i}[n] = x\big(LTn + \tau_i\big), \quad n = 0, 1, \cdots \tag{60}$$

The average sampling rate of the p-channel is p times that of one channel. Note that when the sampling rate of the p-channel is greater than the Nyquist rate, we only need to operate the Nyquist frequency to sample, the actual sampling rate is in the form of

$$\min(\frac{p}{LT}, f_{\mathrm{nyq}}) \tag{61}$$

The theoretical lower bound of the sampling rate is given in [17], which is directly determined by the true bandwidth of the signal:

$$\min(2\lambda(\mathcal{T}), f_{\mathrm{nyq}}) \tag{62}$$

In most cases, the subband bandwidth $\lambda(\mathcal{T})$ and the actual sampling rate does not exceed f_{nyq} (when not satisfied, the sampling rate is f_{nyq}). To ensure reconstruction performance, the parameter p is often not set too low (for instance, it's often chosen to be at least twice the sparsity level $\mathrm{supp}(\mathbf{X})$). It will be seen from above that the theoretical lower bound on the sampling rate is achieved only when $p/LT = pf_s \leq 2\lambda(\mathcal{T}) = 2N_{sig}B$. In other words, when $p = 2\mathrm{supp}(\mathbf{X})$ for the worst case of p, the condition for the actual sampling rate to meet the theoretical lower bound is $K \leq \frac{N_{sig}B}{f_s}$.

C. Proof of Theorem 2

Consider all blocks $\left\{\mathbf{X}^{U_1}, \cdots, \mathbf{X}^{U_M}\right\}$ in $\mathbf{X}$, where there are consecutive corresponding frequency points of length B. For the case $B > f_s$ and each block occupies M rows in $\mathbf{X}$, the length of the frequency point in each $\bar{\mathbf{X}}^{U_i}$ meets

$$l = B - (M-2)f_s < f_s \tag{63}$$

Because $l < f_s$, we know that the non-zero elements of an any sub-block (i.e. $\bar{\mathbf{X}}^{U_i}$) are distributed on both sides and do not intersect on the columns. Conside one block $\bar{\mathbf{X}}^{U_i}$, We let $r \to \infty$ and observe the change in $\bar{\mathbf{X}}^{U_i}$, $\bar{\mathbf{X}}^{U_i}$ gradually changes from an MMV form to an SMV form. In SMV, the sparsity of the signal is determined by the column in $\bar{\mathbf{X}}^{U_i}$ where it is located. It is observed that

$$lim_{r\to\infty,i,j}\mathrm{supp}(\bar{\mathbf{X}}^{U_i}_{:,j}) \leq 1. \tag{64}$$

Thus, the sparsity of each column in $\bar{\mathbf{X}}^{U_i}$ is less than N_{sig}.

A more complex situation is when r is a finite value, assuming $r = r^\star$ is a finite value. In this case, the length of frequency point in each sub-matrix is $\frac{f_s}{r^\star}$. We use reduction to absurdity to prove the condition that the sparsity of each sub-matrix is less than N_{sig}. Assuming that there exists a sub-matrix $\bar{\mathbf{X}}_{S_i}$ with $\mathrm{supp}(\bar{\mathbf{X}}_{S_i}) > N_{sig}$. Also, the non-zero elements of an any sub-block of $\bar{\mathbf{X}}_{S_i}$ do not intersect on the columns. $\bar{\mathbf{X}}_{S_i}$ must contain both non-zero elements on both sides of one block. From (63), we know that the length of any block of $\bar{\mathbf{X}}$ is less than f_s. We can draw a conclusion that

$$\frac{f_s}{r^\star} > f_s - l = (M-1)f_s - B \tag{65}$$

As can be seen, there exist a contradiction between (65) and Theorem 2 , so the length of any column-partition sub-matrix non-zero elements in $\bar{\mathbf{X}}$ must be less or equal than $f_s - l$, which is equivalent to r satisfying

$$r \geq \lceil \frac{f_s}{(M-1)f_s - B} \rceil \tag{66}$$